



The effects of smoking and drinking on the oral and esophageal microbiota of healthy people

Zhengqi Li^{1^}, Yong Liu¹, Lizhou Dou¹, Yueming Zhang¹, Shun He¹, Deli Zhao², Wei Zhang³, Guiqi Wang¹

¹Department of Endoscopy, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China; ²Cancer Center, Feicheng People's Hospital, Feicheng, China; ³Department of Endoscopy, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital & Shenzhen Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen, China

Contributions: (I) Conception and design: G Wang, Z Li; (II) Administrative support: G Wang; (III) Provision of study materials or patients: Y Zhang, S He; (IV) Collection and assembly of data: Y Liu, L Dou, D Zhao, W Zhang; (V) Data analysis and interpretation: Y Liu, Z Li; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Guiqi Wang, MD. Department of Endoscopy, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China. Email: wangguiq@126.com.

Background: To explore the effects of smoking and drinking on the microbiota in the saliva and three segments of the esophagus (upper, middle, and lower) in healthy individuals.

Methods: Paired saliva and brush specimens were obtained from 76 participants who underwent upper gastrointestinal (UGI) endoscopic examination for UGI cancer screening. The esophageal microbiota was investigated by 16S rRNA gene profiling via next-generation sequencing.

Results: The saliva samples from non-smoking and non-drinking participants had a greater abundance of *Neisseria*, *Prevotella*, *Porphyromonas*, and *Rotbia*, and lower levels of *Streptococcus*, *Actinobacillus*, and *Haemophilus* compared to the esophagus. There were no significant differences in the abundance of most bacterial genera in the upper, middle, and lower oesophagus. Similarly, in the saliva of patients who smoke and drink, there was a higher prevalence of *Neisseria*, *Prevotella*, *Porphyromonas*, *Fusobacterium*, and *Rotbia*, and a lower prevalence of *Streptococcus*, *Actinobacillus*, and *Haemophilus* compared to the esophagus. There were no significant differences in the abundance of most genera in the upper, middle, and lower esophagus of patients with a history of drinking and smoking. There were slight differences in the microbiota between smoking and drinking individuals and non-smoking and non-drinking individuals.

Conclusions: This pilot study demonstrated microbial diversity at different taxonomic levels in the oral cavity and esophagus of non-drinking and non-smoking individuals, as well as healthy people who drink and smoke. There was a slight difference in the microbiota between non-drinking and non-smoking people and individuals with a history of drinking and smoking. These results suggested that oral or esophageal cancer caused by smoking and drinking may not be mediated by mechanisms that affect surface microorganisms.

Keywords: Smoking and drinking; oral and oesophageal; microbiota

Submitted Jun 02, 2021. Accepted for publication Jul 15, 2021.

doi: 10.21037/atm-21-3264

View this article at: <https://dx.doi.org/10.21037/atm-21-3264>

[^] ORCID: 0000-0002-1351-8410.

Introduction

The human microbiome of the upper digestive tract is a complex and highly diverse ecosystem (1). The internal balance of microbial diversity is critical to the development of mucus barrier function and immune response to pathogen invasion (2,3). With the development of metagenomics, recent researches have shown that esophageal microbiota, especially the esophageal flora and periodontal pathogens, may be involved in the occurrence and development of oral and esophageal diseases, but the specific mechanisms of action are still unclear (4-6).

As the entry point to the digestive tract, the oral cavity is constantly exposed to a range of inhaled and ingested microorganisms, including more than 700 different types of bacteria, some of which may be related to cancer and other systemic diseases (7,8). The esophagus is a tubular muscle structure approximately 20–27 cm long that connects the mouth and the stomach. In recent years, many studies have reported the composition of the esophageal microbiota of normal healthy people with varying results. In addition, some reports have described the relationship between the oral microbiota (1) and the microbiota of the upper, middle, and lower esophagus in the healthy population. However, studies involving larger sample sizes are required to further understand the esophageal microenvironment (9,10).

Smoking and drinking are important risk factors for esophageal cancer, and much research has focussed on the possible mechanisms (11-14). Esophageal cancer is the eighth most common malignant tumour worldwide, and has the sixth highest mortality of all malignancies (13,15). China has an especially high incidence of esophageal cancer (16), with Feicheng in the Shandong Province being a high-incidence area for esophageal squamous cell cancer (ESCC).

This current study examined the microbiota in the saliva and the upper, middle, and lower oesophageal brush specimens obtained from healthy people in Feicheng. Participants were also given a questionnaire survey regarding their living habits. The composition of the microbial environment was compared between participants with a history of smoking and drinking and participants who did not smoke or drink. This information will provide a greater understanding of the possible role of the gastrointestinal microflora in the development of esophageal cancer in patients who drink and smoke. We present the following article in accordance with the MDAR

and STROBE reporting checklists (available at <https://dx.doi.org/10.21037/atm-21-3264>).

Methods

Study participants

This trial was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Ethics Committee of the National Cancer Center/Cancer Hospital, Chinese Medical College and Peking Union Medical College (No. 17-124/1380). All participants in this study were clearly notified and signed written informed consent forms were obtained.

A total of 76 participants who underwent endoscopic screening conducted by the Chinese Upper Gastrointestinal (UGI) Cancer Project at the Feicheng People's Hospital in Shandong Province, China, between November 2019 and December 2019 were enrolled in this study. All patients were aged 50–70 years and were confirmed by gastroscopy to have no esophageal cancer, low-grade dysplasia (LGD), high-grade dysplasia (HGD), nor ESCC. Patients were excluded from this study if they presented with other systemic diseases; had taken antibiotics, proton pump inhibitors, probiotics, or other preparations within the past month; or presented with oral ulcers in the past month.

Experienced staff collated basic patient characteristics including age, gender, smoking and drinking status, history of medication for digestive system diseases, and oral conditions in the past month.

Sample collection

Prior to endoscopy for UGI cancer screening, patients were requested to avoid any food and drink, and refrain from brushing their teeth. A questionnaire was administered by staff, and a 5 mL saliva sample was collected using saliva collection tubes (SAL2000L, Zeesan, Xiamen, China).

Brush samples of the upper esophagus were collected 18–20 cm from the central incisor, and samples of the middle esophagus and lower esophagus were collected 25–28 cm from the central incisor and 2 cm above the gastroesophageal junction, respectively. Brush specimens were collected using sterile brushes, and the brush head was removed and placed into a sterile tube (Cryovial, 3.0 mL cryogenic tube). All paired specimens were stored at –80 °C immediately after sampling and transported to the

laboratory on dry ice.

DNA extraction, amplification, and sequencing

Total genomic DNA from the saliva and esophageal brush specimens was extracted by applying the CTAB (cetyltrimethylammonium bromide) method. DNA concentration and purity were assessed by running the samples on 1% agarose gels. DNA samples were then diluted to 1 ng/ μ L with sterile water. The universal bacterial primer set was used to amplify the V4 region of the 16S ribosomal RNA (rRNA) gene to generate amplicons. DNA amplification covering the variable V4 region was performed with the 515 F and 806 R primers, incorporating barcode sequences for the 16S rRNA gene. All polymerase chain reactions (PCRs) were conducted in 30 μ L reactions with 15 μ L of Phusion[®] High-Fidelity PCR Master Mix (New England Biolabs), 0.2 μ M forward and reverse primers, and approximately 10 ng of template DNA.

Thermal cycling was performed with 1 minute primary denaturation at 98 °C, followed by 30 cycles of denaturation at 98 °C for 10 seconds, annealing at 50 °C for 30 seconds, 3 second elongation at 72 °C, and 5 minutes elongation at 72 °C. The same volume of 1 \times loading buffer (containing SYBR Green) was mixed with the PCR products, and electrophoresis was performed on a 2% agarose gel for detection. PCR products were mixed in equidensity ratios. A GeneJET[™] Gel Extraction Kit (Thermo Scientific) was used to purify the mixture of PCR products. An Ion Plus Fragment Library Kit (48 reactions; Thermo Scientific) was used to generate sequencing libraries according to the manufacturer's recommendations. Library quality was assessed on a Qubit[®] 2.0 Fluorometer (Thermo Scientific). Finally, the library was sequenced on an Ion S5[™] XL platform, and 600 bp single-end reads were generated.

Sequence processing and taxonomic classification

The Quantitative Insights into Microbial Ecology 2 (QIIME2, <https://qiime2.org/>) platform was used to process the specimens. The UCHIME algorithm (UCHIME algorithm, http://www.drive5.com/usearch/manual/uchime_algo.html) (17) was adopted to compare the reads with the reference database (18) (Silva database, <https://www.arb-silva.de/>) to test chimera sequences, which were subsequently removed (19) and clean reads were obtained. Uparse software (Uparse v7.0.1001, <http://drive5.com/uparse/>) (20) was used to perform sequence analysis.

Sequences with a similarity of $\geq 97\%$ were assigned to the same operational taxonomic unit (OTU). Representative sequences for each OTU were screened for further annotation. For every representative sequence, the Silva database (<https://www.arb-silva.de/>) (18) was applied on the basis of the Mothur algorithm for the annotation of taxonomic information. For research on the phylogenetic association of various OTUs and the difference in the dominant species in different samples (groups), the MUSCLE software (v3.8.31, <http://www.drive5.com/muscle/>) (21) was used to conduct multiple sequence alignments. A standard sequence number corresponding to the sample with the fewest sequences was adopted to normalize OTU abundance information. Subsequent analyses of alpha diversity and beta diversity were all conducted on the basis of these output normalized data.

Statistical analysis

The Student's *t*-tests and chi-squared tests were used to compare demographic data and other characteristics between the non-smoking/non-drinking population and the smoking/drinking population. The software program R Studio (v3.6.1; R Foundation for Statistical Computing, Vienna, Austria) was used to perform all statistical analyses. Alpha diversity differences were determined by using the Wilcoxon rank-sum test. Principal coordinates analysis (PCoA) was performed with the WGCNA package, stat package, and ggplot2 package in R software to identify discrepancies among the independent beta diversity matrices. A distance matrix of unweighted UniFrac distances among previously obtained samples was transformed to a new set of orthogonal axes, by which the maximum variation factor was demonstrated. High-relative-abundance (≥ 0.01) genera were compared between the normal and three other participant groups using the Wilcoxon rank-sum test.

Microbes associated with tumour status were identified by performing linear discriminant analysis effect size (LEfSe) (22). Microbiota constituents with a linear discriminant analysis (LDA) score greater than 4.0 were identified.

Results

Baseline characteristics of the participants

The 76 patients included in this study were divided into two groups according to smoking and alcohol consumption,

Table 1 Baseline characteristics of the participants

Variables	Non-smoking/non-drinking group (n=60)	Smoking/drinking group (n=16)	P
Age, years ($\bar{x}\pm s$)	58.00 \pm 6.07	58.75 \pm 4.81	0.649
Gender, n (%)			0.001
Male	29 (48.33)	15 (93.75)	
Female	31 (51.67)	1 (6.25)	
BMI ($\bar{x}\pm s$)	23.15 \pm 3.26	24.32 \pm 3.12	0.832

BMI, body mass index.

namely the non-smoking/non-drinking group (60 patients) and the smoking/drinking group (16 patients). There were no significant differences in age or body mass index (BMI) between the groups. The smoking/drinking group had more males compared to the non-smoking/non-drinking group ($P=0.001$; *Table 1*).

Differences in the microbial composition at different sample locations in the non-smoking/non-drinking population

Specimens were collected from 4 different locations in the non-smoking/non-drinking participants, namely, the saliva (NSa), the upper esophagus (NUEa), the middle esophagus (NMEa), and the lower esophagus (NLEa).

Figure 1A-1C shows the compositions of the microbiota in the specimens from the 4 different locations. As shown in *Figure 1D,1E*, there were significant differences in the Chao1 indices between NSa and each of the esophageal sites, NUEa, NMEa, and NLEa ($P<0.001$), and significant differences between NUEa and NLEa ($P<0.05$), but no significant differences in the Shannon index. Assessment of beta diversity revealed significant clustering for the unweighted UniFrac distance between NSa and NUEa, NSa and NMEa, NSa and NLEa, NUEa and NMEa, NUEa and NLEa, and NMEa and NLEa ($P<0.01$, *Figure 1F*).

At the phylum level, the abundance of *Firmicutes* in the esophagus was greater than that in the saliva, and the abundance of *Firmicutes* tended to decrease from the upper esophagus to the lower esophagus. There were greater numbers of *Bacteroidota*, *Fusobacteriota*, *Actinobacteriota*, *Gracilibacteria*, and *Cyanobacteria* in the saliva compared to the esophagus. There were no significant differences in most phyla between the upper, middle, and lower esophagus (*Figure 2A*).

At the genera level, *Neisseria*, *Prevotella*, *Porphyromonas*, and *Rotbia* were more abundant in the saliva compared to the esophagus, while *Streptococcus*, *Actinobacillus*, and *Haemophilus* were less abundant in the saliva compared to the esophagus. There were no significant differences in most genera between the upper, middle, and lower esophagus (*Figure 2B*).

LEfSe showed that in the NSa group, the top 3 LDA scores were obtained for *Bacteroidetes*, *Bacteroidota*, and *Bacteroides*. In the NUEa group, the top 3 LDA scores were observed for *Bacilli*, *Firmicutes*, and *Lactobacillales*. In the NMEa group, the top 3 LDA scores were observed with *Pasteurellaceae*, *Pasteurellales*, and *Actinobacillus*. The characteristic class in the NLEa group was *Alphaproteobacteria* (*Figure 2C*).

Differences in the microbial composition at different sample locations in the smoking/drinking population

The specimens from the different sample locations obtained from people who smoke and consume alcohol were designated NSb for saliva, NUEb for samples of the upper esophagus, NMEb for samples of the middle esophagus, and NLEb for samples of the lower esophagus.

Figure 3A-3C shows the composition of the microbiota in the 4 sample locations. As shown in *Figure 3D,3E*, there were significant differences in the Chao1 indices between NSb and NMEb ($P<0.001$), and between NSb and NLEb ($P<0.05$), but no significant differences in the Shannon index. Regarding beta diversity, significant clustering was detected for the unweighted UniFrac distance between NSb and NUEb, NSb and NMEb, NSb and NLEb, NUEb and NMEb, and NMEb and NLEb ($P<0.01$; *Figure 3F*).

At the phylum level, *Firmicutes* was more abundant in the esophagus than in the saliva, and the abundance of *Firmicutes* decreased from the upper esophagus to the lower

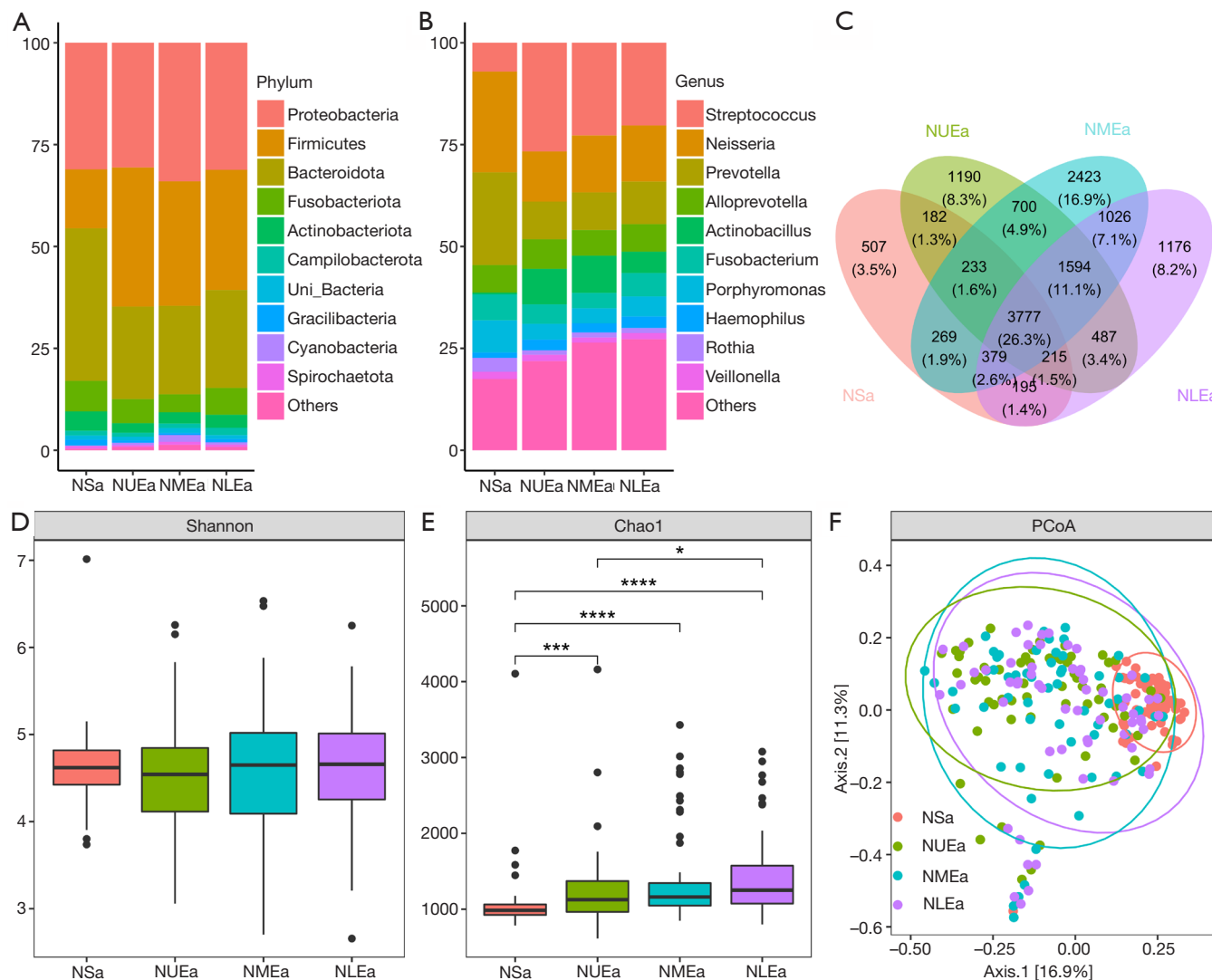


Figure 1 The microbiota characteristics, α diversities, and β diversities in the samples from the non-smoking/non-drinking population. (A) The relative abundances of phyla in the NSa, NUEa, NMEa, and NLEa groups. (B) The relative abundances of genera in the NSa, NUEa, NMEa, and NLEa groups. (C) The observed OTUs in the NSa, NUEa, NMEa, and NLEa groups. (D) The Shannon indices for the NSa, NUEa, NMEa, and NLEa groups. (E) The Chao1 indices for the NSa, NUEa, NMEa, and NLEa groups. (F) The unweighted UniFrac distance ($P < 0.05$) was compared to the PCoA for the NSa, NUEa, NMEa, and NLEa groups. *, $P < 0.05$; ***, $P < 0.001$; ****, $P < 0.0001$. NSa, NUEa, NMEa, and NLEa refer to samples collected from the non-smoking/non-drinking population. NSa, saliva; NUEa, upper esophagus; NMEa, middle esophagus; NLEa, lower esophagus; OTU, operational taxonomic unit; PCoA, principal coordinates analysis.

esophagus. *Bacteroidota*, *Fusobacteriota*, *Gracilibacteria*, and *Acidobacteriota* were more abundant in the saliva than in the esophagus. There were no significant differences in most phyla between the upper, middle, and lower oesophagus (Figure 4A).

In terms of genera, there was a higher abundance of *Neisseria*, *Prevotella*, *Porphyromonas*, *Fusobacterium* and *Rothia*

in the saliva compared to the esophagus, while *Streptococcus*, *Actinobacillus*, and *Haemophilus* were less abundant in the saliva compared to the esophagus. There were no significant differences in most genera between the upper, middle, and lower esophagus (Figure 4B).

LEfSe demonstrated that in the NSb group, the phylum with the top 3 LDA scores were *Bacteroidetes*, *Bacteroides*,

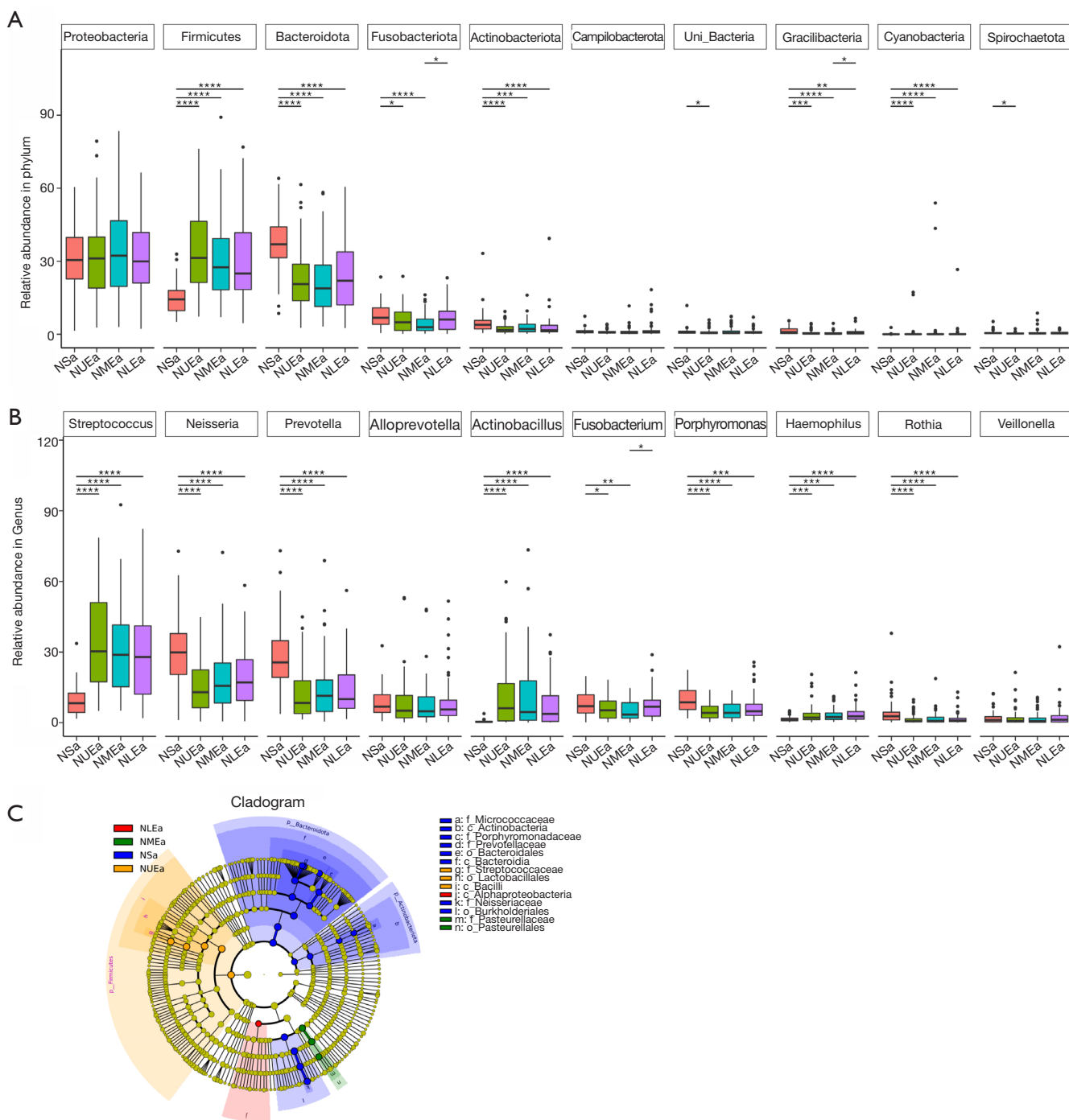


Figure 2 The microbiota characteristics of the samples from the non-smoking/non-drinking population. (A) Mean relative abundances of the top 10 most plentiful microbial phyla among the NSa, NUEa, NMEa, and NLEa groups. (B) Mean relative abundances of the top 10 most common genera of microbes among the NSa, NUEa, NMEa, and NLEa groups. (C) Circular cladogram for niche specialization of microbial compositions among the NSa, NUEa, NMEa, and NLEa groups using linear discriminant analysis effect size (LEfSe) analysis of the abundance patterns of bacterial taxa. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$. NSa, NUEa, NMEa, and NLEa refer to the non-smoking/non-drinking population. NSa, saliva; NUEa, upper esophagus; NMEa, middle esophagus; NLEa, lower esophagus.

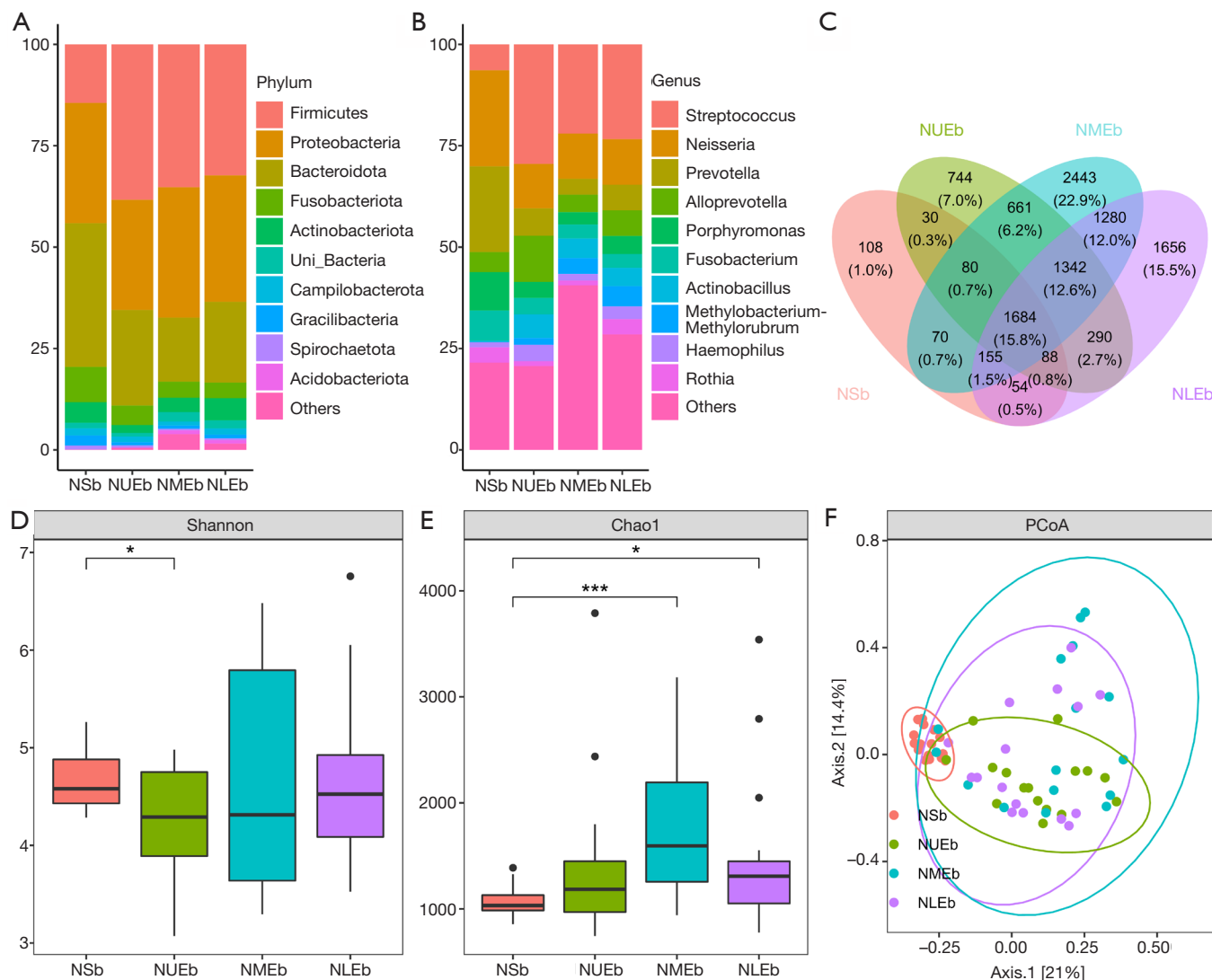


Figure 3 The microbiota characteristics, α diversities, and β diversities in the samples from the smoking/drinking population. (A) The relative abundances of phyla in the NSb, NUEb, NMEb, and NLEb groups. (B) The relative abundances of genera in the NSb, NUEb, NMEb, and NLEb groups. (C) The observed OTUs in the NSb, NUEb, NMEb, and NLEb groups. (D) The Shannon indices for the NSb, NUEb, NMEb, and NLEb groups. (E) The Chao1 indices for the NSb, NUEb, NMEb, and NLEb groups. (F) The unweighted UniFrac distance ($P < 0.05$) was compared to the PCoA for the NSb, NUEb, NMEb and NLEb groups. *, $P < 0.05$; ***, $P < 0.001$. NSb, NUEb, NMEb, and NLEb refer to the smoking/drinking population. NSb, saliva; NUEb, upper esophagus; NMEb, middle esophagus; NLEb, lower esophagus; OTU, operational taxonomic unit; PCoA, principal coordinates analysis.

and *Bacteroides*. In the NUEb group, the top 3 LDA scores were obtained for *Bacilli*, *Firmicutes*, and *Lactobacillales*. In the NMEa group, the top 1 LDA score was obtained for *Alphaproteobacteria*. The characteristic taxa in NLEb group were *Rhizobiales*, *Beijerinckiaceae*, and *Methylobacterium_Methylorubrum* (Figure 4C).

Microbiota associations and differences between paired non-smoking/non-drinking samples and smoking/drinking samples from the same location

As shown in Figures 5-8, there were no significant microbiota differences in the saliva samples, upper esophageal samples, nor the lower esophageal samples

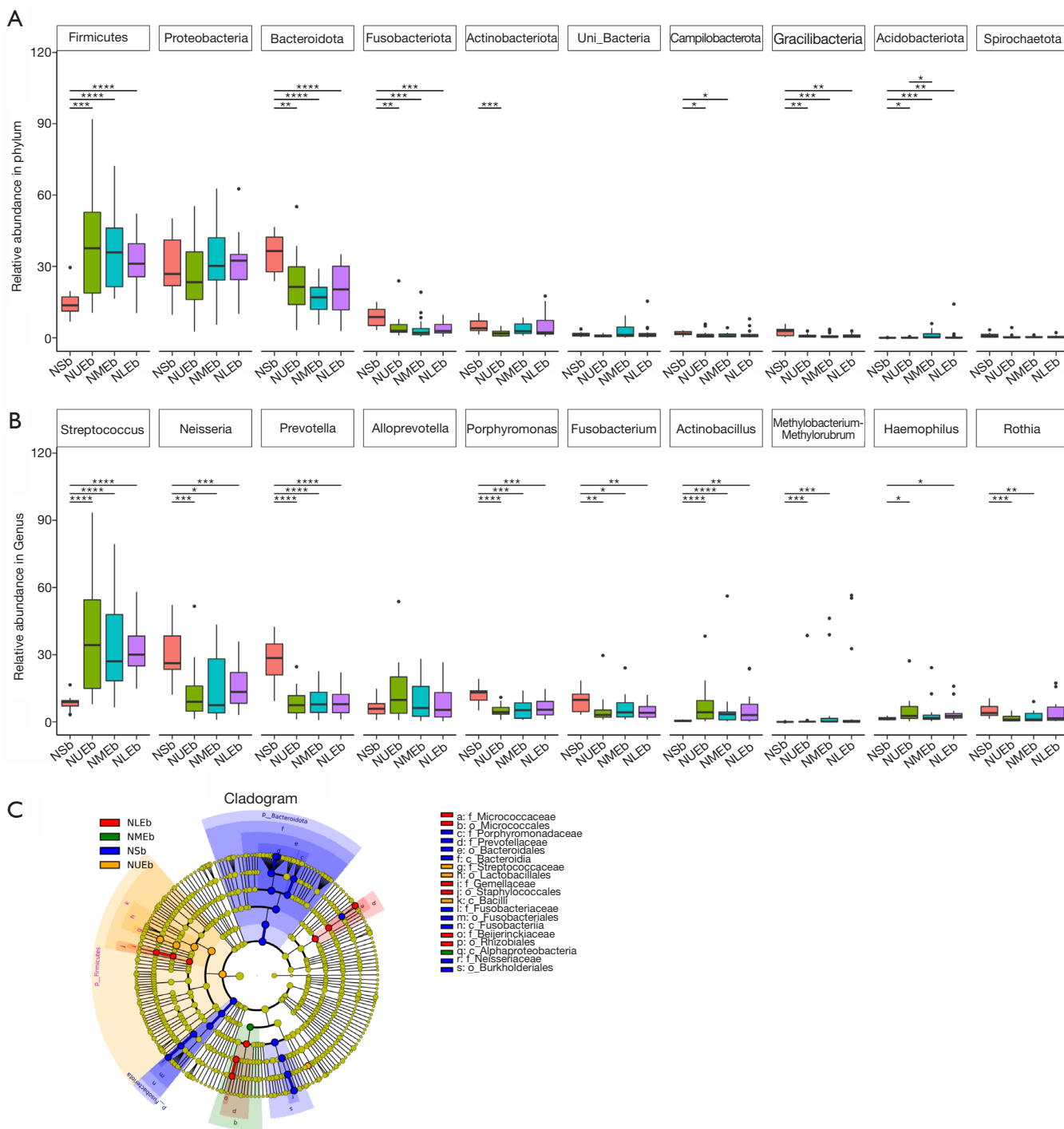


Figure 4 The microbiota characteristics of the samples from the smoking/drinking population. of relative abundance in the NSb, NUEb, NMEb and NLEb groups. (A) Mean relative abundances of the top 10 most plentiful microbial phyla among the NSb, NUEb, NMEb, and NLEb groups. (B) Mean relative abundances of the top 10 most common genera of microbes among the NSa, NSb, NUEb, NMEb, and NLEb groups. (C) Circular cladogram for the niche specialization of microbial compositions among the NSb, NUEb, NMEb, and NLEb groups using linear discriminant analysis effect size (LEfSe) analysis of the abundance patterns of bacterial taxa. *, P<0.05; **, P<0.01; ***, P<0.001; ****, P<0.0001. NSb, NUEb, NMEb, and NLEb refer to the smoking/drinking population. NSb, saliva; NUEb, upper esophagus; NMEb, middle esophagus; NLEb, lower esophagus.

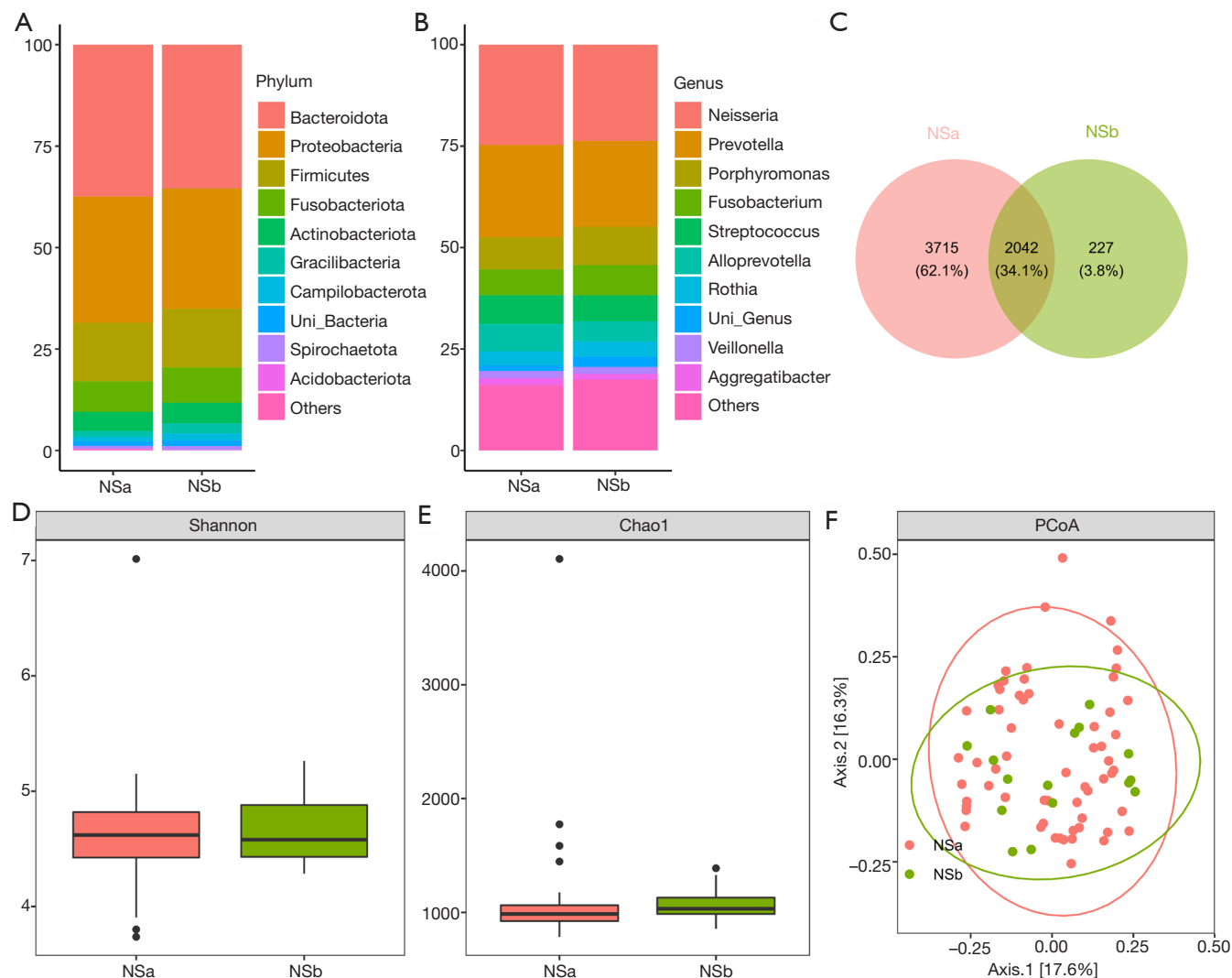


Figure 5 The microbiota characteristics, α diversities, and β diversities in the saliva samples from the non-smoking/non-drinking population and the smoking/drinking population. (A) The relative abundances of phyla in the NSa and NSb groups. (B) The relative abundances of genera in the NSa and NSb groups. (C) The observed OTUs in the NSa and NSb groups. (D) The Shannon indices for the NSa and NSb groups. (E) The Chao1 indices for the NSa and NSb groups. (F) The unweighted UniFrac distance ($P < 0.05$) was compared to the PCoA for the NSa and NSb groups. NSa, saliva sample from non-smoking/non-drinking patients; NSb, saliva sample from smoking/drinking patients; OTU, operational taxonomic unit; PCoA, principal coordinates analysis.

between the non-smoking/non-drinking group and the smoking/drinking group. However, there was a significant difference in the Chao1 index between the middle esophageal samples from the two groups ($P < 0.05$; Figure 7E).

At the phylum level, *Campylobacterota* and *Gracilibacteria* were significantly more abundant in the saliva samples from the smoking/drinking group compared to the non-smoking/non-drinking group. *Cyanobacteria* was also significantly

more abundant in the middle esophagus of the smoking/drinking group compared to the non-smoking/non-drinking group. There were no significant differences between the two groups in the upper esophagus and the lower esophagus (Figure 9).

At the genus level, *Uni_Genus* was significantly more abundant in the saliva samples from the smoking/drinking group compared to the non-smoking/non-drinking group. Similarly, *Bacteroides* and *Uni_Genus* were significantly

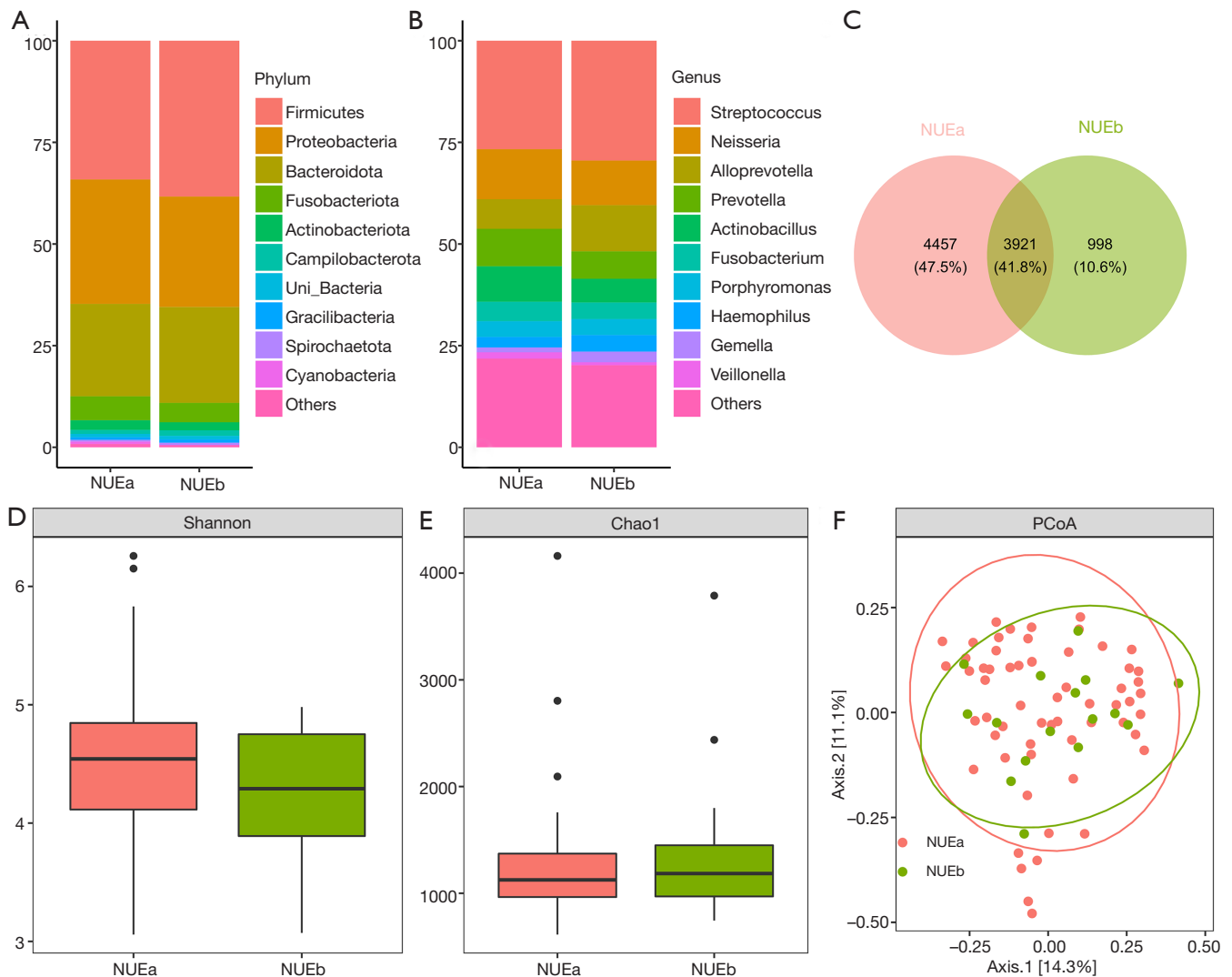


Figure 6 The microbiota characteristics, α diversities, and β diversities in the upper esophageal samples from the non-smoking/non-drinking population and the smoking/drinking population. (A) The relative abundances of phyla in the NUEa and NUEb groups. (B) The relative abundances of genera in the NUEa and NUEb groups. (C) The observed OTUs in the NUEa and NUEb groups. (D) The Shannon indices for the NUEa and NUEb groups. (E) The Chao1 indices for the NUEa and NUEb groups. (F) The unweighted UniFrac distance ($P < 0.05$) was compared to the PCoA for the NUEa and NUEb groups. NUEa, upper esophageal sample from non-smoking/non-drinking patients; NUEb, upper esophageal sample from smoking/drinking patients; OTU, operational taxonomic unit; PCoA, principal coordinates analysis.

more abundant in the middle esophagus of the smoking/drinking group compared to the non-smoking/non-drinking group. *Gemella* was significantly less abundant in the upper esophagus of non-smoking/non-drinking patients compared to drinking/smoking patients. Similarly, *Rotbia* was significantly less abundant in the lower esophagus of the non-smoking/non-drinking group compared to the smoking/drinking group (Figure 10).

Discussion

Smoking and drinking are considered significant risk factors for the onset of esophageal cancer and oropharyngeal cancer. However, the precise carcinogenic mechanisms have not yet been fully elucidated, and may involve a complex process of multi-stage and multi-factor interactions (11,12,14,23). Indeed, changes in the digestive tract flora caused by

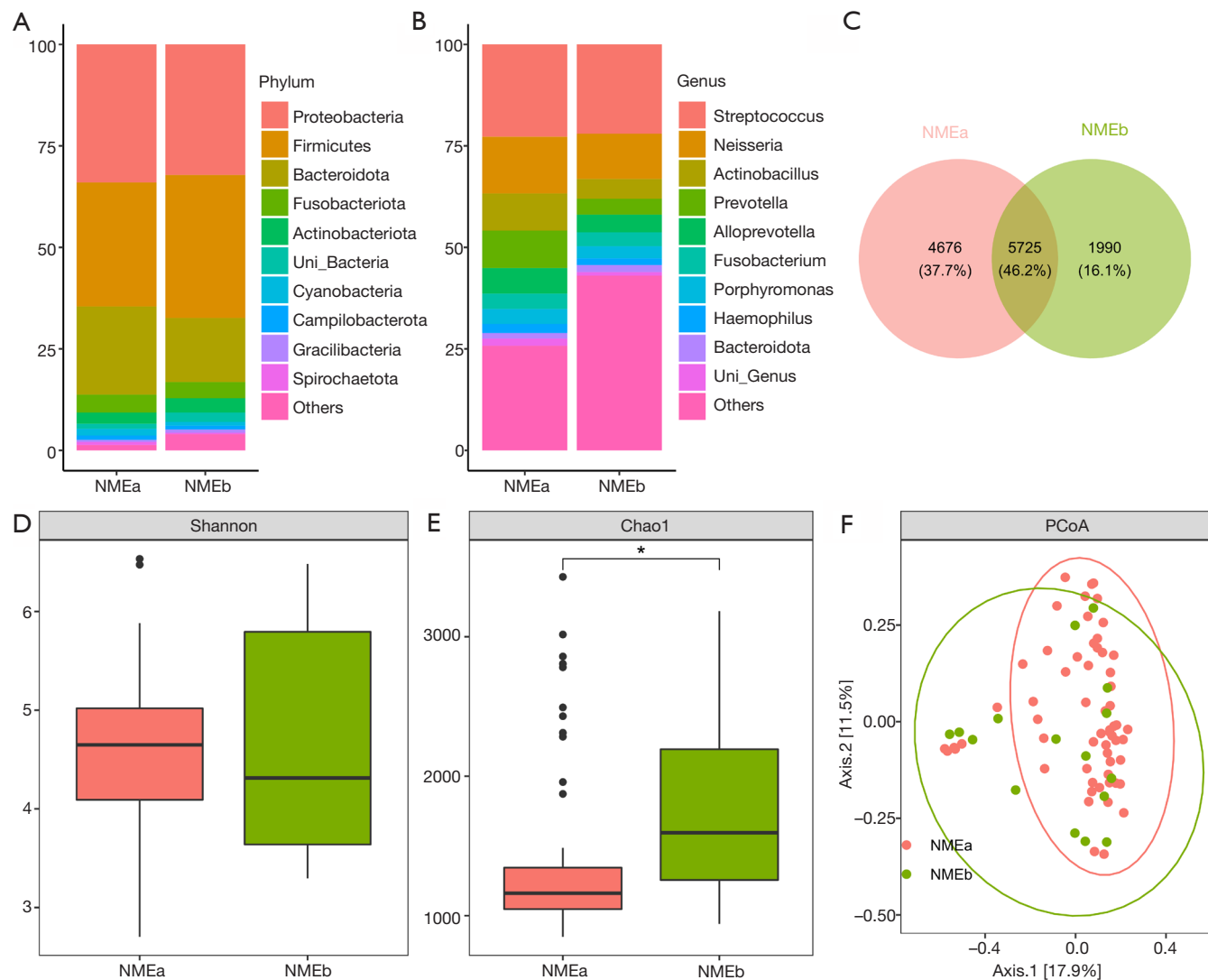


Figure 7 The microbiota characteristics, α diversities, and β diversities in the middle esophageal samples from the non-smoking/non-drinking population and the smoking/drinking population. (A) The relative abundances of phyla in the NMEa and NMEb groups. (B) The relative abundances of genera in the NMEa and NMEb groups. (C) The observed OTUs in the NMEa and NMEb groups. (D) The Shannon indices for the NMEa and NMEb groups. (E) The Chao1 indices for the NMEa and NMEb groups. (F) The unweighted UniFrac distance ($P < 0.05$) was compared to the PCoA for the NMEa and NMEb groups. *, $P < 0.05$. NMEa, middle esophageal sample from non-smoking/non-drinking patients; NMEb, middle esophageal sample from smoking/drinking patients; OTU, operational taxonomic unit; PCoA, principal coordinates analysis.

smoking and drinking may be related to the occurrence and development of cancer. In healthy individuals, microbiome balance is considered to be dynamic because it varies with endogenous and exogenous factors. Human lifestyles and experiences can rapidly and profoundly alter the stability of host-associated microbial communities. Oral microorganisms may cause system diseases by activating the

immune pathways in the body. However, the use of specific oral microorganisms as biomarkers for disease diagnosis and treatment is still lacking in sufficient prospective studies (24). Therefore, studying the occurrence and progression of microbial changes in the oral cavity and esophagus of the smoking/drinking population is of great significance for the early diagnosis, comprehensive

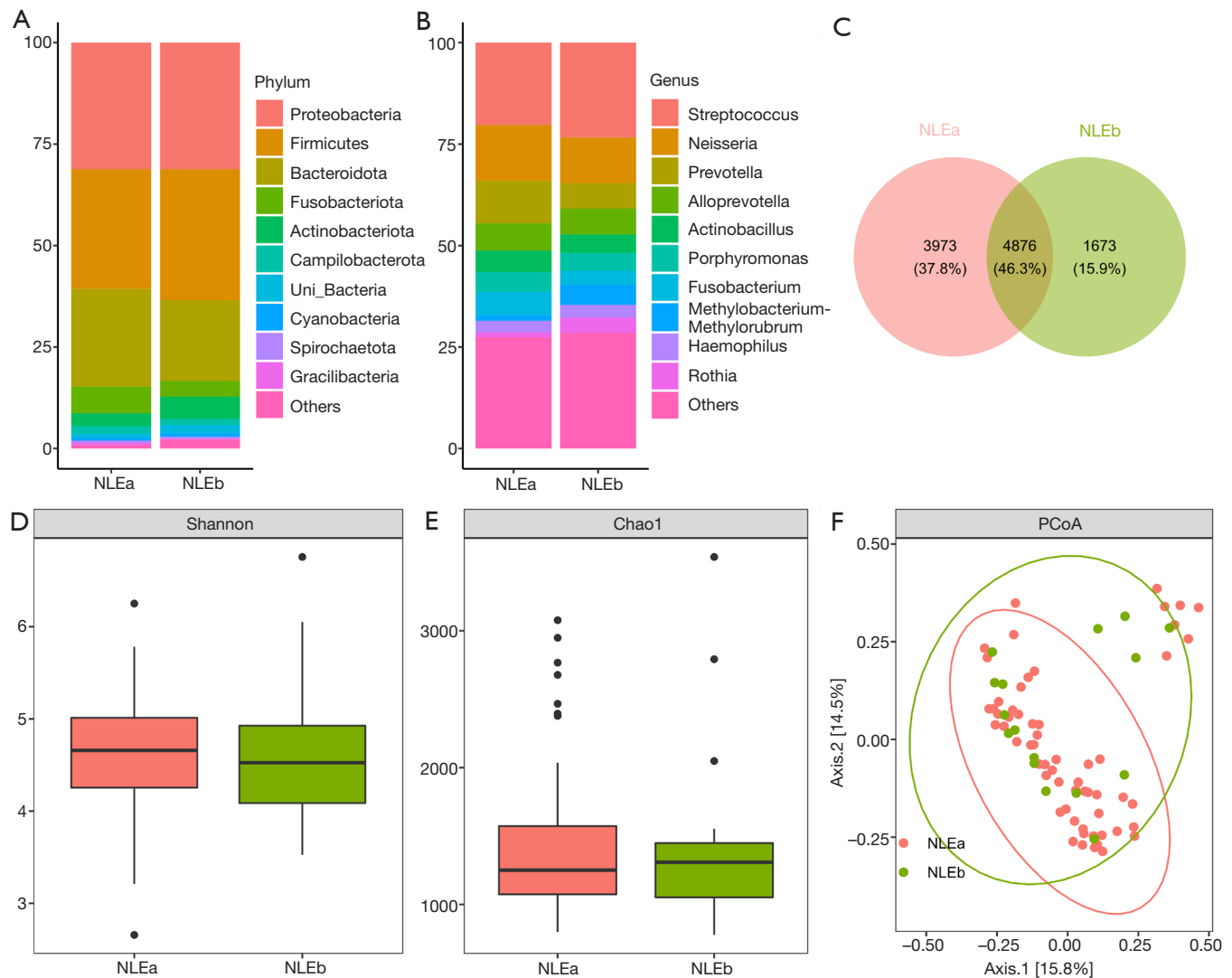


Figure 8 The microbiota characteristics, α diversities, and β diversities in the lower esophageal samples from the non-smoking/non-drinking population and the smoking/drinking population. (A) The relative abundances of phyla in the NLEa and NLEb groups. (B) The relative abundances of genera in the NLEa and NLEb groups. (C) The observed OTUs in the NLEa and NLEb groups. (D) The Shannon indices for the NLEa and NLEb groups. (E) The Chao1 indices for the NLEa and NLEb groups. (F) The unweighted UniFrac distance ($P < 0.05$) was compared to the PCoA for the NLEa and NLEb groups. NLEa, lower esophageal sample from non-smoking/non-drinking patients; NLEb, lower esophageal sample from smoking/drinking patients; OTU, operational taxonomic unit; PCoA, principal coordinates analysis.

treatment, prognosis evaluation, and treatment efficacy of oropharyngeal and esophageal cancers.

This study conducted in a healthy population from a high-incidence region for esophageal cancer in China, demonstrated the contiguity and preference of the microbiota in the oral cavity and three segments of the esophagus using 16S rRNA gene sequencing technology. The five most abundant phyla in the oral cavity and upper,

middle, and lower sites of the esophagus of the non-smoking/non-drinking population were *Proteobacteria*, *Firmicutes*, *Bacteroidetes*, *Actinobacteria*, and *Fusobacteria*, which was consistent with previous studies (1,25). Interestingly, the five most abundant phyla in the oral cavity and upper, middle, and lower sites of the esophagus in people who smoke and drink were identical to that observed in non-smoking/non-drinking people. However, the most

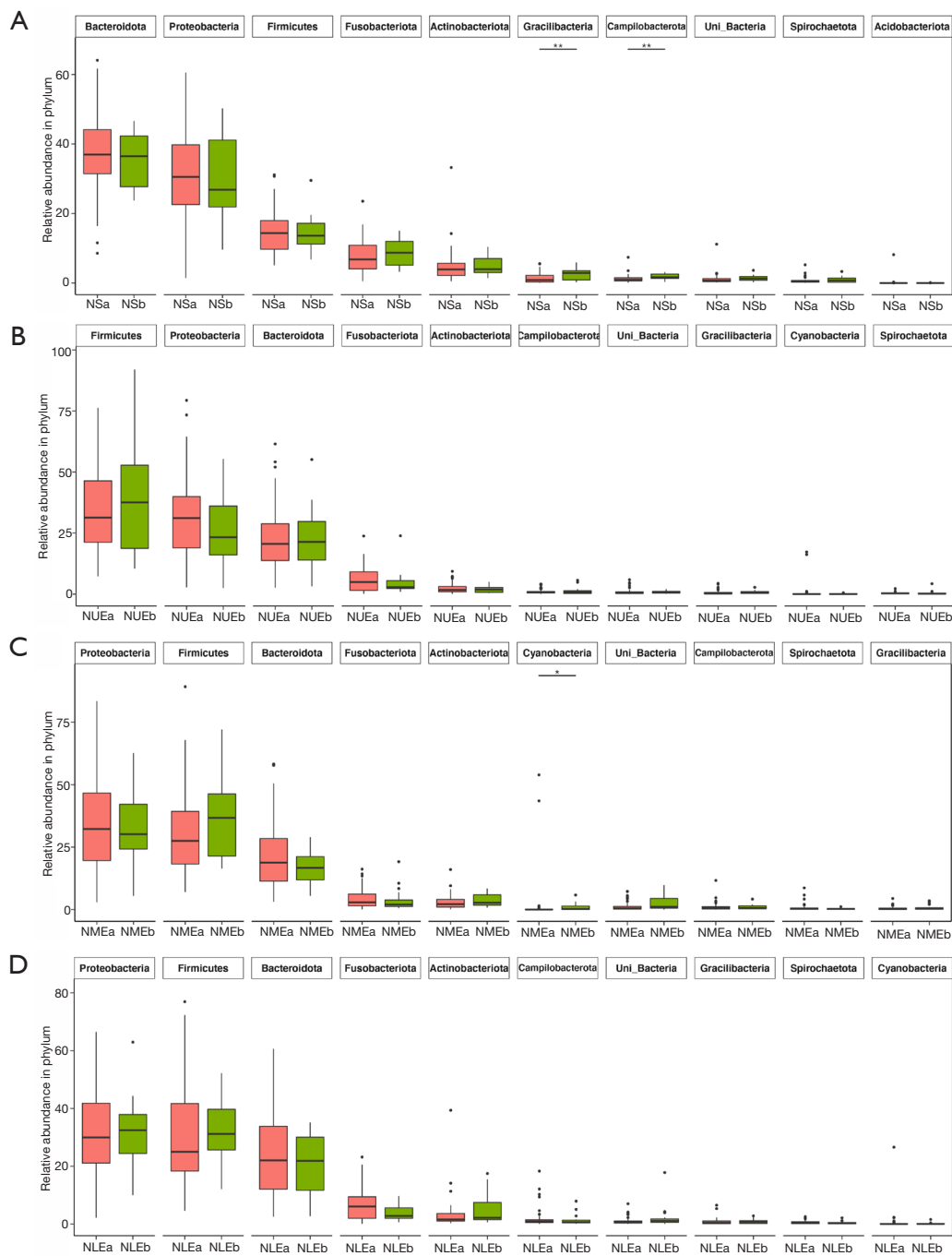


Figure 9 Mean relative abundances of the top 10 most copious microbial phyla between paired non-smoking/non-drinking samples and smoking/drinking samples from the same location. (A) Mean relative abundances of the top 10 most plentiful microbial phyla between the NSa and NSb groups. (B) Mean relative abundances of the top 10 most abundant microbial phyla between the NUEa and NUEb groups. (C) Mean relative abundances of the top 10 most abundant microbial phyla between the NMEa and NMEb groups. (D) Mean relative abundances of the top 10 most abundant microbial phyla between the NLEa and NLEb groups. *, $P < 0.05$; **, $P < 0.01$. NSa, saliva sample from non-smoking/non-drinking patients; NSb, saliva sample from smoking/drinking patients; NUEa, upper esophageal sample from non-smoking/non-drinking patients; NUEb, upper esophageal sample from smoking/drinking patients; NMEa, middle esophageal sample from non-smoking/non-drinking patients; NMEb, middle esophageal sample from smoking/drinking patients; NLEa, lower esophageal sample from non-smoking/non-drinking patients; NLEb, lower esophageal sample from smoking/drinking patients.

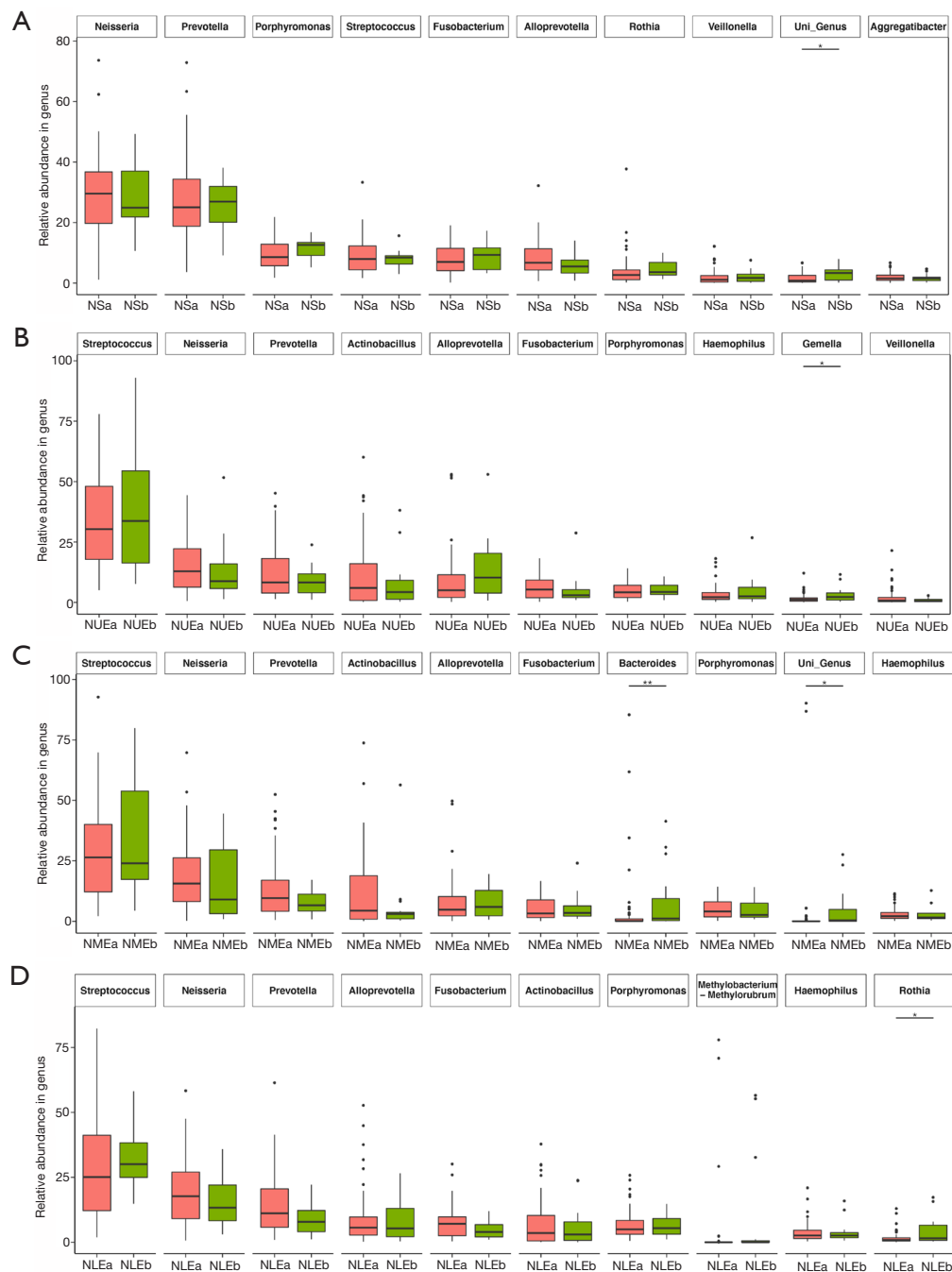


Figure 10 Mean relative abundances of the top 10 most abundant microbial genera between paired non-smoking/non-drinking samples and smoking/drinking samples from the same location. (A) Mean relative abundances of the top 10 most abundant microbial genera between the NSa and NSb groups. (B) Mean relative abundances of the top 10 most abundant microbial genera between the NUEa and NUEb groups. (C) Mean relative abundances of the top 10 most abundant microbial genera between the NMEa and NMEb groups. (D) Mean relative abundances of the top 10 most abundant microbial genera between the NLEa and NLEb groups. *, $P < 0.05$; **, $P < 0.01$. NSa, saliva sample from non-smoking/non-drinking patients; NSb, saliva sample from smoking/drinking patients; NUEa, upper esophageal sample from non-smoking/non-drinking patients; NUEb, upper esophageal sample from smoking/drinking patients; NMEa, middle esophageal sample from non-smoking/non-drinking patients; NMEb, middle esophageal sample from smoking/drinking patients; NLEa, lower esophageal sample from non-smoking/non-drinking patients; NLEb, lower esophageal sample from smoking/drinking patients.

abundant phylum in non-smoking/non-drinking people was *Proteobacteria*, while in people who smoke and drink, it was *Firmicutes*. In the non-smoking/non-drinking population, the Chao1 richness analyses showed that the microbial profiles in the specimens from the saliva were statistically less diverse than those in the specimens from the three sites of the oesophagus. Again, this agreed with reports by Dong and colleagues (1). Furthermore, the lower esophageal samples were more diverse than those from the upper oesophagus. However, in the smoking/drinking population, there were no significant differences in the microbial diversity between the saliva and upper esophagus samples.

The abundance of bacteria at the phylum and the genus levels was different among the saliva and the three sites in the oesophagus, and LEfSe showed that the characteristic species at different sampling sites were different from those reported by Dong *et al.* (1). To our knowledge, this is the first study to document the microbiological differences between the saliva and the three sites of the esophagus in the smoking/drinking population, and these differences warrant further investigation.

Studies had shown that cigarette smoking and drinking alcohol can cause oral microbiota imbalance, and it was speculated that oral microbes may play a role in smoking-related or alcohol-related diseases, but the specific mechanism was not clear (26,27). This study also compared the changes in the oral and esophageal microbiotas between the smoking/drinking population and the non-smoking/non-drinking population. In the saliva samples, there was a greater abundance of *Gracilibacteria* and *Campylobacterota* in the smoking/drinking group compared to the non-smoking/non-drinking group. *Streptococcus sp.*, *Peptostreptococcus sp.*, *Prevotella sp.*, *Porphyromonas gingivalis*, and *Capnocytophaga gingivalis* are often observed in oral squamous cell carcinomas (8,28). However, in this current study, no changes in the abundance of these bacteria were observed. In the upper esophageal samples, the levels of *Gemella* in the smoking/drinking group were higher than those in the non-smoking/non-drinking group. In the middle esophageal samples, the abundance of *Bacteroides* in the smoking/drinking group was higher than that in the non-smoking/non-drinking group. In the lower esophageal samples, the abundance of *Rothia* was higher in the smoking/drinking group compared to the non-smoking/non-drinking group. Interestingly, *Veillonella*, *Prevotella*, *Haemophilus*, *Neisseria*, *Granulicatella*, and *Fusobacterium* are all thought to be associated with esophageal adenocarcinomas (29), however, we did not detect any significant changes in the abundance

of these bacteria in this study. Furthermore, high levels of *Haemophilus*, *Neisseria*, and *Porphyromonas* are thought to be associated with esophageal squamous cell carcinomas (30,31), however, no differences in the abundance of these bacteria were observed between the smoking/drinking population and the non-smoking/non-drinking population in this study. Therefore, according to our research, smoking and drinking had a slight impact on the microbial composition in the oral cavity and esophagus in normal healthy people, however, smoking and drinking did not appear to have any significant effect on the population of oral or esophageal cancer surface-dominant bacteria. This finding suggested that oral or esophageal cancer due to smoking and drinking may not be mediated by mechanisms that affect surface microorganisms. This research provided a plot study for further researches of the microbial basis of non-communicable diseases.

There were several limitations to this study. First, the sample size was small, especially the number of people who smoke and drink. Second, the proportion of males in the smoking/drinking population was higher than that in the non-smoking/non-drinking population, and this may have created bias in the study results.

Conclusions

This research has demonstrated the microbial diversity at different taxonomic levels in the oral cavity and esophagus of non-smoking/non-drinking individuals and drinking/smoking healthy people. However, there were no significant differences in the population of oral or esophageal cancer surface-dominant bacteria between the two groups. These results suggested that oral or esophageal cancer due to smoking and drinking may not be mediated by mechanisms that affect surface microorganisms.

Acknowledgments

Funding: This research was supported by grants from the National Key Research and Development Program of China (Grant No. 2018YFC1313100 and 2016YFC1302800); the Shenzhen Sanming Medical Project (No. SZSM201911008); the CAMS Innovation Fund for Medical Sciences (CIFMS; Approval no. 2016-I2M-1-001, 2017-I2M-1-001, and 2019-I2M-2-004); the PUMC Youth Fund and the Fundamental Research Funds for the Central Universities (Approval No. 2017320012); the Capital Health Improvement and Research Fund (Grant No. 2020-2-4025); and the PUMC

Graduate Innovation Fund (Approval No. 2019-1002-81).

Footnote

Reporting Checklist: The authors have completed the MDAR and STROBE reporting checklists. Available at <https://dx.doi.org/10.21037/atm-21-3264>

Data Sharing Statement: Available at <https://dx.doi.org/10.21037/atm-21-3264>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/atm-21-3264>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This trial was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Ethics Committee of the National Cancer Center/Cancer Hospital, Chinese Medical College and Peking Union Medical College (No. 17-124/1380), and informed consent was obtained from all individual participants.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Dong L, Yin J, Zhao J, et al. Microbial Similarity and Preference for Specific Sites in Healthy Oral Cavity and Esophagus. *Front Microbiol* 2018;9:1603.
- Segata N, Haake SK, Mannon P, et al. Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples. *Genome Biol* 2012;13:R42.
- Zhang X, Pan Z. Influence of microbiota on immunity and immunotherapy for gastric and esophageal cancers. *Gastroenterol Rep (Oxf)* 2020;8:206-14.
- Li L, Li C, Lv M, et al. Correlation between alterations of gut microbiota and miR-122-5p expression in patients with type 2 diabetes mellitus. *Ann Transl Med* 2020;8:1481.
- Di Pilato V, Freschi G, Ringressi MN, et al. The esophageal microbiota in health and disease. *Ann N Y Acad Sci* 2016;1381:21-33.
- Okereke I, Hamilton C, Wenzholz A, et al. Associations of the microbiome and esophageal disease. *J Thorac Dis* 2019;11:S1588-S1593.
- Zhang Y, Niu Q, Fan W, et al. Oral microbiota and gastrointestinal cancer. *Onco Targets Ther* 2019;12:4721-8.
- Tuominen H, Rautava J. Oral Microbiota and Cancer Development. *Pathobiology* 2021;88:116-26.
- Coker OO, Dai Z, Nie Y, et al. Mucosal microbiome dysbiosis in gastric carcinogenesis. *GUT* 2018;67:1024-32.
- Fillon SA, Harris JK, Wagner BD, et al. Novel device to sample the esophageal microbiome--the esophageal string test. *PLoS One* 2012;7:e42938.
- Batra R, Malhotra GK, Singh S, et al. Managing Squamous Cell Esophageal Cancer. *Surg Clin North Am* 2019;99:529-41.
- Thrumurthy SG, Chaudry MA, Thrumurthy SSD, et al. Oesophageal cancer: risks, prevention, and diagnosis. *BMJ* 2019;366:l4373.
- Abnet CC, Arnold M, Wei W. Epidemiology of Esophageal Squamous Cell Carcinoma. *Gastroenterology* 2018;154:360-73.
- Matejcic M, Gunter MJ, Ferrari P. Alcohol metabolism and oesophageal cancer: a systematic review of the evidence. *Carcinogenesis* 2017;38:859-72.
- Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015;136:E359-E386.
- Qiu ML, Lin JB, Li X, et al. Current state of esophageal cancer surgery in China: a national database analysis. *BMC Cancer* 2019;19:1064.
- Edgar RC, Haas BJ, Clemente JC, et al. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 2011;27:2194-200.
- Quast C, Pruesse E, Yilmaz P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2013;41:D590-D596.
- Haas BJ, Gevers D, Earl AM, et al. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res*

- 2011;21:494-504.
20. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* 2013;10:996-8.
 21. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32:1792-7.
 22. Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. *Genome Biol* 2011;12:R60.
 23. Langille MG, Zaneveld J, Caporaso JG, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* 2013;31:814-21.
 24. Jia G, Zhi A, Lai PFH, et al. The oral microbiota – a mechanistic role for systemic diseases. *Brit Dent J* 2018;224:447-55.
 25. Gormley M, Dudding T, Sanderson E, et al. A multivariable Mendelian randomization analysis investigating smoking and alcohol consumption in oral and oropharyngeal cancer. *Nat Commun* 2020;11:6071.
 26. Fan X, Peters BA, Jacobs EJ, et al. Drinking alcohol is associated with variation in the human oral microbiome in a large study of American adults. *Microbiome* 2018;6:59.
 27. Rodríguez-Rabassa M, López P, Rodríguez-Santiago R, et al. Cigarette Smoking Modulation of Saliva Microbial Composition and Cytokine Levels. *Int J Environ Res Public Health* 2018;15:2479.
 28. Karpiński TM. Role of Oral Microbiota in Cancer Development. *Microorganisms* 2019;7:20.
 29. Lv J, Guo L, Liu J, et al. Alteration of the esophageal microbiota in Barrett's esophagus and esophageal adenocarcinoma. *World J Gastroenterol* 2019;25:2149-61.
 30. Li D, He R, Hou G, et al. Characterization of the Esophageal Microbiota and Prediction of the Metabolic Pathways Involved in Esophageal Cancer. *Front Cell Infect Microbiol* 2020;10:268.
 31. Li M, Shao D, Zhou J, et al. Signatures within esophageal microbiota with progression of esophageal squamous cell carcinoma. *Chin J Cancer Res* 2020;32:755-67.

(English Language Editor: J. Teoh)

Cite this article as: Li Z, Liu Y, Dou L, Zhang Y, He S, Zhao D, Zhang W, Wang G. The effects of smoking and drinking on the oral and esophageal microbiota of healthy people. *Ann Transl Med* 2021;9(15):1244. doi: 10.21037/atm-21-3264